Chapter 8

Fastest Mixing and Vertex Expansion

We study a very recent result by Oleskar-Taylor and Zanetti [OZ21] relating the fastest mixing time to the vertex expansion of a graph, proving a surprising and beautiful Cheeger inequality for vertex expansion.

The fastest mixing time problem was proposed by Boyd, Diaconis and Xiao [BDX04]. In this problem, we are given an undirected graph $G = (V, E)$ and a target probability distribution $\pi : V \to \mathbb{R}$. The task is to assign a transition probability $P(u, v)$ on each edge $uv \in E(G)$, so that the stationary distribution of random walks with transition matrix P is π. The objective is to find a transition matrix P that minimizes the mixing time to π, among all transition matrices with stationary distribution π. We know from chapter 6 that the mixing time to the stationary distribution is approximately inversely proportional to the spectral gap $1 - \alpha_2(P)$ of the transition matrix P, where $1 = \alpha_1(P) \geq \alpha_2(P) \geq \cdots \geq \alpha_{|V|}(P) \geq -1$ are the eigenvalues of P. The fastest mixing time problem is thus formulated in [BDX04] by the maximum spectral gap achievable through a "reweighting" P of the adjacency matrix of G.

Definition 8.1 (Maximum Reweighted Spectral Gap [BDX04])**.** *Given an undirected graph $G = (V, E)$ and a probability distribution π on V, the maximum reweighted spectral gap is defined as*

$$\begin{aligned} \lambda_2^*(G) \;:=\; \max_{P \geq 0} \quad & 1 - \alpha_2(P) & \\ \text{subject to} \quad & P(u,v) = 0 & \forall uv \notin E \\ & \sum_{v \in V} P(u,v) = 1 & \forall u \in V \\ & \pi(u)P(u,v) = \pi(v)P(v,u) & \forall uv \in E. \end{aligned}$$

The last constraint is called the time reversible condition, which is to ensure that the stationary distribution of P is π. Note that $\lambda_2^*(G) = \max_{P \geq 0}(1 - \alpha_2(P)) = \max_{P \geq 0} \lambda_2(I - P)$, which is the maximum reweighted second smallest eigenvalue of the normalized Laplacian matrix of G subject to the above constraints.

Boyd, Diaconis and Xiao showed that this optimization problem can be written as a semidefinite program and thus $\lambda_2^*(G)$ can be computed in polynomial time. Subsequently, the fastest mixing time problem has been studied in various work (see the references in [OZ21]), but no general characterization was known. Roch [Roc05] showed that the vertex expansion $\psi(G)$ defined in Definition 7.8 is an upper bound on the optimal spectral gap $\lambda_2^*(G)$. Very recently, Olesker-Taylor and Zanetti [OZ21] proved that small vertex expansion is qualitatively the only obstruction for fastest mixing time to the uniform distribution.

Theorem 8.2 (Cheeger Inequality for Vertex Expansion [OZ21]). *For any undirected graph $G = (V, E)$ and the uniform distribution $\pi = \vec{1}/|V|$,*

$$\frac{\psi(G)^2}{\log |V|} \lesssim \lambda_2^*(G) \lesssim \psi(G).$$

In terms of the fastest mixing time $\tau^*(G)$ to the uniform distribution,

$$\frac{1}{\psi(G)} \lesssim \tau^*(G) \lesssim \frac{\log^2 |V|}{\psi^2(G)}.$$

Note the analogy to the Cheeger's inequality in Theorem 4.3, where spectral gap is replaced by maximum reweighted spectral gap and edge conductance is replaced by vertex expansion.

Unlike Cheeger's inequality for edge conductance where $\phi(G)^2 \lesssim \lambda_2(G) \lesssim \phi(G)$, it is noted in [OZ21] that the $\log |V|$ term may not be completely removed: Louis, Raghavendra and Vempala [LRV13] proved that there is no polynomial time algorithm that can distinguish between $\psi(G) \leq \epsilon$ and $\psi(G) \gtrsim \sqrt{\epsilon \log d}$ for every $\epsilon > 0$ where d is the maximum degree of the graph G, assuming the small-set expansion conjecture of Raghavendra and Steurer [RS10]. So, if the $\log |V|$ factor in Theorem 8.2 can be completely removed, then $\lambda_2^*(G)$ is a polynomial time computable quantity that can be used to distinguish between the two cases, disproving the small-set expansion conjecture.

Remark 8.3 (Uniform Distribution and Self-Loops). *We will make two assumptions about the problem. One is that the target distribution is the uniform distribution. Another is that the graph has a self-loop on each vertex, so that the problem in Definition 8.1 is always feasible. In the context of Markov chains, this corresponds to allowing a non-negative holding probability on each vertex.*

8.1 Dual Program for Fastest Mixing

To prove Theorem 8.2, Oleskar-Taylor and Zanetti use a dual minimization program obtained by Roch [Roc05] of the primal maximization program in Definition 8.1, to relate $\lambda_2^*(G)$ to the minimum vertex expansion of the input graph. We will use Von Neumann's minimax theorem to derive Roch's dual program.

Theorem 8.4 (Von Neumann's Minimax Theorem). *Let X, Y be compact convex sets. If f is a real-valued continuous function on $X \times Y$ with $f(x, \cdot)$ concave on Y for all $x \in X$ and $f(\cdot, y)$ convex on X for all $y \in Y$, then*

$$\min_{x \in X} \max_{y \in Y} f(x, y) = \max_{y \in Y} \min_{x \in X} f(x, y).$$

In the proof, we will use the following semidefinite program for computing the second eigenvalue, which is an extension of the spectral program in Lemma 4.4 to higher dimension but with the same optimal value.

Lemma 8.5 (Semidefinite Program for the Second Eigenvalue). *Let $P \in \mathbb{R}^{n \times n}$ be a reweighted matrix of a graph $G = (V, E)$ satisfying the constraints in Definition 8.1. Then*

$$1 - \alpha_2(P) = \min_{f: V \to \mathbb{R}^n,\ \sum_{v \in V} f(v) = 0} \frac{\sum_{uv \in E} \| f(u) - f(v) \|^2 \cdot P(u, v)}{\sum_{v \in V} \| f(v) \|^2}.$$

Proof. As explained in Definition 8.1, $1-\alpha_2(P) = \lambda_2(I-P)$ where $I-P$ is the normalized Laplacian matrix of the weighted graph P with weighted degree one for each vertex. By Lemma 4.4,

$$1 - \alpha_2(P) = \min_{f:V\to\mathbb{R},\ \sum_{v\in V} f(v)=0} \frac{\sum_{uv\in E} |f(u)-f(v)|^2 \cdot P(u,v)}{\sum_{v\in V} f(v)^2},$$

which is almost the same as in the statement, except that $f : V \to \mathbb{R}$ instead of $f : V \to \mathbb{R}^n$ as in the statement. Clearly, by considering all $f : V \to \mathbb{R}^n$, the feasible set could only be bigger and so the optimal value could only be smaller. On the other hand, given a solution $f : V \to \mathbb{R}^n$, by using the inequality

$$\min_{1\le i\le n} \frac{a_i}{b_i} \le \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}$$

on the coordinates of $f : V \to \mathbb{R}^n$, we see that the best coordinate gives a one-dimensional solution $f : V \to \mathbb{R}$ with objective value as good as that of the n-dimensional solution $f : V \to \mathbb{R}^n$. To summarize, the relaxation from $f : V \to \mathbb{R}$ to $f : V \to \mathbb{R}^n$ is an exact relaxation. □

To see that it is a semidefinite program, recall that a positive semidefinite matrix Y can be written as $F^T F$ where $F \in \mathbb{R}^{n\times n}$ by Fact 2.7. We associate each column v of F to $f(v)$, so that $Y_{u,v} = \langle f(u), f(v)\rangle$ for all $u, v \in V$. Then the above program can be rewritten as

$$\begin{aligned} \min \quad & \sum_{uv\in E} (Y_{u,u} - 2Y_{u,v} + Y_{v,v}) \cdot P(u,v) \\ \text{subject to} \quad & \sum_{v\in V} Y_{v,v} = 1 \\ & \sum_{u,v\in V} Y_{u,v} = 0 \\ & Y \succcurlyeq 0, \end{aligned}$$

where the objective function is the numerator in Lemma 8.5, the first constraint is normalizing the denominator in Lemma 8.5 to one, the second constraint is equivalent to the constraint $\sum_{v\in V} f(v) = 0$, and the last constraint is to ensure the correspondence $Y_{u,v} = \langle f(u), f(v)\rangle$ for all $u, v \in V$. So, the program in Lemma 8.5 can be written as optimizing a linear function with linear constraints on the entries of a positive semidefinite matrix Y, and this is a semidefinite program that can be solved in polynomial time.

The reason that we use the above semidefinite program for the second eigenvalue instead of the spectral program is that the set of feasible solutions is a convex set (while it is not the case for the spectral program), and this would allow us to apply the Von-Neumann minimax theorem to derive the following dual program by Roch.

Proposition 8.6 (Dual Program for Fastest Mixing [Roc05, OZ21]). *Given an undirected graph $G = (V,E)$ with a self-loop on each vertex and the uniform distribution $\pi = \vec{1}/|V|$ on V, the following semidefinite program is dual to the primal program in Definition 8.1 with strong duality*

$\lambda_2^*(G) = \gamma(G)$ *where*

$$\begin{aligned} \gamma(G) := \min_{f:V\to\mathbb{R}^n,\ g:V\to\mathbb{R}_{\geq 0}} \quad & \sum_{v\in V} g(v) \\ \text{subject to} \quad & \sum_{v\in V} \|f(v)\|^2 = 1 \\ & \sum_{v\in V} f(v) = \vec{0} \\ & g(u) + g(v) \geq \|f(u) - f(v)\|^2 \qquad \forall uv \in E. \end{aligned}$$

Proof. For a fixed P, by Lemma 8.5,

$$1 - \alpha_2(P) = \min_{f:V\to\mathbb{R}^n,\ \sum_{v\in V} f(v)=0} \frac{\sum_{uv\in E} \|f(u)-f(v)\|^2 \cdot P(u,v)}{\sum_{v\in V}\|f(v)\|^2}.$$

The maximum reweighted spectral gap in Definition 8.1 can thus be formulated as

$$\begin{aligned} \lambda_2^*(G) = \max_{P\geq 0}(1-\alpha_2(P)) \quad = \quad & \max_{P\geq 0} \min_{f:V\to\mathbb{R}^n,\ \sum_{v\in V} f(v)=0} \frac{\sum_{uv\in E}\|f(u)-f(v)\|^2 \cdot P(u,v)}{\sum_{v\in V}\|f(v)\|^2} \\ \text{subject to} \quad & P(u,v) = 0 \quad \forall uv \notin E \\ & \sum_{v\in V} P(u,v) = 1 \quad \forall u \in V \\ & P = P^T. \end{aligned}$$

Check that the assumptions in the Von Neumann minimax Theorem 8.4 are satisfied, and so we can switch the order of the max and the min and obtain the dual program

$$\gamma(G) \quad := \quad \min_{f:V\to\mathbb{R}^n,\ \sum_{v\in V} f(v)=0} \ \max_{P\geq 0} \frac{\sum_{uv\in E}\|f(u)-f(v)\|^2 \cdot P(u,v)}{\sum_{v\in V}\|f(v)\|^2},$$

subjected to the same constraints on P as above.

For a fixed $f : V \to \mathbb{R}^n$, note that the inner maximization problem is a linear program over the entries of P, and so we can reformuate it using LP duality to obtain

$$\begin{aligned} \gamma(G) \quad = \quad & \min_{f:V\to\mathbb{R}^n,\ \sum_{v\in V} f(v)=0} \ \min_{g\geq 0} \sum_{v\in V} g(v) \\ \text{subject to} \quad & g(u)+g(v) \geq \frac{\|f(u)-f(v)\|^2}{\sum_{v\in V}\|f(v)\|^2} \quad \forall uv \in E, \end{aligned}$$

where $g(u)$ is a dual variable for the constraint $\sum_{v\in V} P(u,v) = 1$. Note that the constraint $g \geq 0$ is from the assumption that there is a self-loop at each vertex. Normalizing so that $\sum_{v\in V}\|f(v)\|^2 = 1$ gives the statement. □

We remark that the self-loop assumption is to ensure that the dual program has the inequality $g \geq 0$. This is a crucial but subtle condition that will be used only once, and we will point it out when it is used.

One-Dimensional Dual Program and Random Projection

The first step in the proof of Theorem 8.2 is to project the solution $f : V \to \mathbb{R}^n$ to $\gamma(G)$ into a 1-dimensional solution $f : V \to \mathbb{R}$ as follows.

Definition 8.7 (One-Dimensional Dual Program for Fastest Mixing [OZ21]). *Given an undirected graph $G = (V, E)$, $\gamma^{(1)}(G)$ is defined to be the program:*

$$\begin{aligned}
\gamma^{(1)}(G) := \min_{f:V\to\mathbb{R},\ g:V\to\mathbb{R}_{\geq 0}} \quad & \sum_{v\in V} g(v) \\
\text{subject to} \quad & \sum_{v\in V} f(v)^2 = 1 \\
& \sum_{v\in V} f(v) = 0 \\
& g(u) + g(v) \geq (f(u) - f(v))^2 \qquad \forall uv \in E.
\end{aligned}$$

A very important result in metric embedding is the dimension reduction theorem by Johnson and Lindenstrauss, which says that n high-dimensional vectors can be projected into $O(\log n)$-dimensional vectors so that the pairwise Euclidean distances are approximately preserved.

Theorem 8.8 (Johnson-Lindenstrauss Lemma). *Given $0 < \epsilon < 1$, a set X of n points in $\mathbb{R}^m$, there is a linear map $A : \mathbb{R}^m \to \mathbb{R}^k$ for $k \lesssim \ln(n)/\epsilon^2$ such that for all $u, v \in X$ it holds that*

$$(1-\epsilon)\|u - v\|_2^2 \leq \|Au - Av\|_2^2 \leq (1+\epsilon)\|u - v\|_2^2.$$

Apply the Johnson-Lindenstrauss lemma to the n-dimensional solution f in Proposition 8.6 to obtain a $O(\log n)$-dimensional solution f' with only constant distortion, and then use the "best" coordinate in f' as a solution to Definition 8.7, one can prove the following bounds between the two programs. Note that the $\log|V|$ factor in Theorem 8.2 is from this dimension reduction step.

Problem 8.9 (Dimension Reduction [OZ21]). *For any undirected graph G,*

$$\gamma(G) \leq \gamma^{(1)}(G) \lesssim \log|V(G)| \cdot \gamma(G).$$

The main step in the proof of Theorem 8.2 is the following Cheeger-type inequality between the 1-dimensional program in Definition 8.7 and the vertex expansion of the graph.

Theorem 8.10 (Cheeger Inequality for Vertex Expansion [OZ21]). *For any undirected graph G,*

$$\psi(G)^2 \lesssim \gamma^{(1)}(G) \lesssim \psi(G).$$

Combining Proposition 8.6 and Problem 8.9 and Theorem 8.10 gives

$$\psi(G)^2 \lesssim \gamma^{(1)}(G) \lesssim \log|V| \cdot \gamma(G) = \log|V| \cdot \lambda_2^*(G) \quad \text{and} \quad \lambda_2^*(G) = \gamma(G) \leq \gamma^{(1)}(G) \lesssim \psi(G),$$

proving Theorem 8.2. Henceforth, our goal is to prove Theorem 8.10, although we will need to do one more transformation described in the next section before getting to the main proof.

8.2 Matching Expansion

Instead of reasoning about the vertex expansion directly, Oleskar-Taylor and Zanetti defined an interesting new concept called the matching expansion, and showed that it is closely related to the vertex expansion and is easier to relate to the 1-dimensional dual program in Definition 8.7.

Definition 8.11 (Matching Expansion [OZ21]). *Let $G = (V, E)$ be an undirected weighted graph with a weight $w(e)$ on each edge $e \in E$. Given a subset of edges $F \subseteq E$, let the weight of a maximum matching in F be*

$$\nu(F) = \max_{\text{matching } M \subseteq F} \sum_{e \in M} w(e).$$

Define the matching expansion of a subset $S \subseteq V$ and of the graph as

$$\psi_\nu(S) = \frac{\nu(\delta(S))}{|S|} \quad \text{and} \quad \psi_\nu(G) = \min_{S: 0 < |S| \leq |V|/2} \psi_\nu(S).$$

Note that while vertex expansion of a set in Definition 7.8 could be much larger than one, the matching expansion of a set is always at most one (in the case when $w(e) = 1$ for all $e \in E$), as is the edge conductance of a set in Definition 4.2. However, it can be shown that the vertex expansion of a graph is about the same as the matching expansion of a graph.

Problem 8.12 (Matching Expansion and Vertex Expansion). *Let G be an undirected graph where every edge is of weight one. Then*

$$\psi_\nu(G) \leq \psi(G) \leq 4\psi_\nu(G).$$

The main technical work in [OZ21] is in proving the following Cheeger inequality for matching expansion.

Theorem 8.13 (Cheeger Inequality for Matching Expansion [OZ21]). *For any undirected graph G where every edge is of weight one,*

$$\psi_\nu(G)^2 \lesssim \gamma^{(1)}(G) \lesssim \psi_\nu(G).$$

It should be clear that Problem 8.12 and Theorem 8.13 imply Theorem 8.10, which in turn implies Theorem 8.2. Our goal is then to prove Theorem 8.13.

Maximum Matching, Auxiliary Directed Graphs, and Directed Matching

The intuition that matching is relevant to the problem is from the constraints in the 1-dimensional dual program in Definition 8.7. The following lemma follows from weak duality of linear programming and is easy to see directly. This is the only place that the constraint $g \geq 0$ is used crucially, so pay attention when the following lemma is used in the main proof.

Lemma 8.14 (Matching and Vertex Cover). *Let $G = (V, E)$ be an undirected graph with a weight $w(e)$ on each edge $e \in E$. The weighted matching number $\nu(E)$ is upper bounded by*

$$\begin{aligned} & \sum_{v \in V} g(v) \\ \text{subject to} \quad & g(u) + g(v) \geq w(u, v) \quad && \forall uv \in E \\ & g(v) \geq 0 && \forall v \in V. \end{aligned}$$

From this perspective, a good way to interpret the solution to Definition 8.7 is that g is a weighted fractional vertex cover when each edge has weight $(f(u) - f(v))^2$. The following auxiliary graphs will be used in the main proof.

Definition 8.15 (Auxiliary Graphs). *Let $G = (V, E)$ be an undirected graph and f, g be a solution to the program in Definition 8.7. Define G_f to be the weighted undirected graph where each edge uv in E has weight $|f(u)^2 - f(v)^2|$. Define $\overrightarrow{G_f}$ to be the orientation of G_f where there is a directed edge uv with weight $f(u)^2 - f(v)^2$ if and only if $uv \in E(G)$ and $f(u) > f(v)$.*

A matching in an undirected graph is a subgraph in which every vertex is of degree at most one. The following analog of directed matching will be used in the proof.

Definition 8.16 (Directed Matching). *Given a directed graph $\overrightarrow{G} = (V, \overrightarrow{E})$ with a weight $w(e)$ on each edge $e \in \overrightarrow{E}$, a subset of edges $\overrightarrow{F} \subseteq \overrightarrow{E}$ is a directed matching if the indegree and the outdegree of each vertex in $\overrightarrow{F}$ is at most one. Let $\nu(\overrightarrow{E})$ be the maximum weight of a directed matching in $\overrightarrow{E}$.*

A simple combinatorial argument shows that the maximum weight of an undirected matching is within a constant factor of the maximum weight of a directed matching.

Exercise 8.17. *For any edge-weighted graph $G = (V, E)$ and any orientation $\overrightarrow{G} = (V, \overrightarrow{E})$,*

$$\nu(E) \leq \nu(\overrightarrow{E}) \leq 4\nu(E).$$

8.3 Cheeger Inequality for Matching Expansion

As in the proof of Cheeger's inequality in Theorem 4.3, one direction is the easy direction where we see that $\gamma^{(1)}(G)$ is a relaxation for the matching expansion $\psi_\nu(G)$, and another direction is the hard direction where we round a fractional solution to $\gamma^{(1)}(G)$ to obtain an integral solution to $\psi_\nu(G)$.

Easy Direction

There are two ways to see the easy direction. One way is to plug in a binary solution defined by a set S minimizing the matching expansion to upper bound $\gamma^{(1)}(G)$.

Proposition 8.18 (Easy Direction for Matching Expansion [OZ21]). *For any undirected graph G where every edge is of weight one,*

$$\gamma^{(1)}(G) \lesssim \psi_\nu(G).$$

Proof. Given $S \subseteq V$, plug in

$$f(v) = \frac{1}{|S|} \frac{\sqrt{|S||V-S|}}{\sqrt{|S| + |V-S|}} \text{ for } v \in S, \quad \text{and } f(v) = -\frac{1}{|V-S|} \frac{\sqrt{|S||V-S|}}{\sqrt{|S| + |V-S|}} \text{ otherwise.}$$

Let M be a maximum matching in $\delta(S)$. Set $g(v) = 2/|S|$ for $v \in M$ and $g(v) = 0$ otherwise. Check that it is a feasible solution to $\gamma^{(1)}(G)$ in Definition 8.7 with objective value at most $4\psi_\nu(S)$. □

Another way is to understand the easy direction of Theorem 8.2 directly, as to use the edge conductance of a reweighted graph H of G to certify the vertex expansion of the input graph G.

Proposition 8.19 (Vertex Expansion through Edge Expansion). *Let H be an edge-reweighted graph of $G = (V, E)$ with weighted adjacency matrix P satisfying the constraints in the primal program in Definition 8.1. Then $\phi(H) \leq \psi(G)$ where $\phi(H)$ is the weighted edge conductance of H.*

Proof. As the reweighted matrix P satisfies the constraints in Definition 8.1, the graph H is a weighted 1-regular graph and so its weighted edge conductance is simply

$$\phi(H) = \min_{S:0<\mathrm{vol}_w(S)\leq\frac{1}{2}\mathrm{vol}_w(V)} \frac{w(\delta(S))}{\mathrm{vol}_w(S)} = \min_{S:0<|S|\leq\frac{1}{2}|V|} \frac{w(\delta(S))}{|S|},$$

where we denote $w(u, v) = P(u, v)$ as the weight of an edge and $w(\delta(S)) = \sum_{e\in\delta(S)} w(e)$. Observe the important point that $|\partial(S)| \geq w(\delta(S))$, because each edge in $\delta(S)$ has an endpoint in $\partial(S)$ and each vertex in $\partial(S)$ has weighted degree one, and so $|\partial(S)| = \sum_{v\in\partial(S)} \deg_w(v) \geq w(\delta(S))$. Therefore,

$$\phi(H) = \min_{S:0<|S|\leq\frac{1}{2}|V|} \frac{w(\delta(S))}{|S|} \leq \min_{S:0<|S|\leq\frac{1}{2}|V|} \frac{|\partial(S)|}{|S|} = \psi(G).$$

□

By Proposition 8.19, the edge conductance of any edge reweighted graph H of G satisfying the constraints in Definition 8.1 is a lower bound on the vertex expansion of G. To prove the best lower bound on the vertex expansion of G, we thus maximum the edge conductance of an edge reweighted graph H. Note that the edge conductance of the reweighted graph H is lower bounded by the spectral gap of the reweighted matrix P by the easy direction of Cheeger's inequality in Theorem 4.3. Therefore,

$$\lambda_2^*(G) = \max_{H:H \text{ is a reweighting of } G} \lambda_2(H) \leq \max_{H:H \text{ is a reweighting of } G} 2\phi(H) \leq 2\psi(G).$$

To summarize, a good way to understand the easy direction of the new Cheeger inequality for vertex expansion in Theorem 8.2 is that it is a way to certify the vertex expansion of a graph through a reduction to the edge conductance and spectral gap of a reweighted graph. Very interestingly, the hard direction proves that there is always a reweighted graph so that this reduction works well to certify the vertex expansion.

Hard Direction

The structure of the hard direction is similar to that for Cheeger's inequality in chapter 4. The first step is a truncation step to ensure that the output set S satisfies $|S| \leq |V|/2$. Again, as in the truncation step in Lemma 4.6 in the hard direction of Cheeger's inequality, the condition $\sum_{v\in V} f(v) = 0$ is used to trade for the non-negativity condition and the support-size condition.

Problem 8.20 (Truncation). *Let $G = (V, E)$ be an undirected graph and $\pi = \vec{1}/|V|$ be the uniform distribution. Given a solution f, g to $\gamma^{(1)}(G)$ in Definition 8.7, there is a solution x, y with $x \geq 0$ and $y \geq 0$ and $|\mathrm{supp}(x)| \leq |V|/2$ such that*

$$\sum_{v\in V} y(v) \lesssim \gamma^{(1)}(G)$$

$$\sum_{v\in V} x(v)^2 = 1$$

$$y(u) + y(v) \geq (x(u) - x(v))^2 \qquad \forall uv \in E.$$

Again, the main step is to apply threshold rounding on the solution in Problem 8.20 to find a set S with small matching expansion.

Proposition 8.21 (Hard Direction for Matching Expansion [OZ21]). *Let $G = (V, E)$ be an undirected graph and $\pi = \vec{1}/|V|$ be the uniform distribution. Given a solution x, y satisfying the conditions in Problem 8.20, there is a set $S \subseteq \mathrm{supp}(x)$ with $\psi_\nu(S) \lesssim \sqrt{\gamma^{(1)}(G)}$.*

Proof. Let $S_t := \{v \in V \mid x(v)^2 > t\}$ be a level set for $t \geq 0$. Choose t uniform randomly, Trevisan's argument implies that

$$\min_t \psi_\nu(S_t) \leq \frac{\int_0^\infty \nu(\delta(S_t))dt}{\int_0^\infty |S_t|dt},$$

so that we can compute the numerator and the denominator separately.

The denominator is

$$\int_0^\infty |S_t|dt = \int_0^\infty \sum_{v \in V} \mathbb{1}(v \in S_t)dt = \sum_{v \in V} \int_0^\infty \mathbb{1}(x(v)^2 > t)dt = \sum_{v \in V} x(v)^2 = 1.$$

To bound the numerator, Oleskar-Taylor and Zanetti consider the auxiliary graphs G_x and $\overrightarrow{G_x}$ in Definition 8.15. Using some combinatorial arguments about matchings, they proved a key lemma in Lemma 8.22 that

$$\int_0^\infty \nu(\delta(S_t))dt \leq 8\nu(G_x).$$

Assuming Lemma 8.22, let M be a maximum weighted matching in G_x, we further bound $\nu(G_x)$ by standard Cauchy-Schwarz manipulation so that

$$\begin{aligned}
\nu(G_x) &= \sum_{uv \in M} \left|x(u)^2 - x(v)^2\right| \\
&= \sum_{uv \in M} \left|x(u) - x(v)\right| \cdot \left|x(u) + x(v)\right| \\
&\leq \sqrt{\sum_{uv \in M} \big(x(u) - x(v)\big)^2} \sqrt{\sum_{uv \in M} \big(x(u) + x(v)\big)^2} \\
&\leq \sqrt{\sum_{uv \in M} \big(x(u) - x(v)\big)^2} \sqrt{\sum_{v \in V} 2x(v)^2} \\
&= \sqrt{2 \sum_{uv \in M} \big(x(u) - x(v)\big)^2}.
\end{aligned}$$

where the last inequality holds because M is a matching so that each vertex is of degree one in M.

Next, we use the weak duality between matching and vertex cover stated in Lemma 8.14 to relate the RHS to the solution y in Problem 8.20, so that

$$\sum_{uv \in M} \big(x(u) - x(v)\big)^2 \leq \sum_{v \in V} y(v) \lesssim \gamma^{(1)}(G),$$

where the first inequality is where the constraint $y \geq 0$ is crucially used. Therefore, we conclude that

$$\min_t \psi_\nu(S_t) \leq \frac{\int_0^\infty \nu(\delta(S_t))dt}{\int_0^\infty |S_t|dt} \leq 8\nu(G_x) \leq 8\sqrt{2 \sum_{uv \in M} \big(x(u) - x(v)\big)^2} \lesssim \sqrt{\gamma^{(1)}(G)}.$$

□

It remains to prove the key lemma about the numerator. Let M_t be a maximum matching in $\delta(S_t)$. Then $\int_0^\infty \nu(\delta(S_t))dt = \int_0^\infty |M_t|dt$. The main idea is to prove that a greedy directed matching $\overrightarrow{M}$ in $\overrightarrow{G_x}$ satisfies $|\overrightarrow{M} \cap \delta(S_t)| \geq \frac{1}{2}|M_t|$ for every t. That is, the fixed matching $\overrightarrow{M}$ is almost as good as the maximum matching in every threshold set S_t.

Lemma 8.22. *Let $G = (V, E)$ be an undirected graph and $\pi = \vec{1}/|V|$ be the uniform distribution. Given a solution x, y satisfying the conditions in Problem 8.20, let $S_t := \{v \in V \mid x(v)^2 > t\}$ for $t \geq 0$, then*

$$\int_0^\infty \nu(\delta(S_t))dt \leq 8\nu(G_x).$$

Proof. Using the definitions of G_x and $\overrightarrow{G_x}$ in Definition 8.15, we will prove that $\int_0^\infty \nu(\delta(S_t))dt \leq 2\nu(\overrightarrow{G_x})$, and then the lemma follows from Exercise 8.17.

To prove $\int_0^\infty \nu(\delta(S_t))dt \leq 2\nu(\overrightarrow{G_x})$, we consider a greedy directed matching $\overrightarrow{M}$ in $\overrightarrow{G_f}$, which is obtained by sorting the directed edges in non-increasing order of weights and greedily adding edges to the directed matching whenever possible.

Let M_t be a maximum matching in $\delta(S_t)$. Note that there could be a different maximum matching for each t. The key observation is that the fixed greedy matching $\overrightarrow{M}$ satisfies $|\overrightarrow{M} \cap \delta(S_t)| \geq |M_t|/2$ for each t. To see this, let uv be an edge in M_t with $x(u) > x(v)$. Suppose $uv \notin \overrightarrow{M}$. Since $\overrightarrow{M}$ is a greedy directed matching, when uv was considered and was not added to $\overrightarrow{M}$, then either u has outdegree one or v has indegree one at that time, as otherwise we could add the edge uv to $\overrightarrow{M}$. In either case, say $uw \in \overrightarrow{M}$, then it must hold that $x(u)^2 - x(w)^2 \geq x(u)^2 - x(v)^2$, as the edges are considered in a non-increasing order of weights. Since uw is at least as long as uv, the edge uw is in every threshold cut that uv is in, and so $uw \in \delta(S_t) \cap \overrightarrow{M}$. Using this argument, we can map each edge $uv \in M_t$ to some other edge in $\delta(S_t) \cap \overrightarrow{M}$ sharing an endpoint with uv. Crucially, since M_t is a matching, each edge in $\overrightarrow{M}$ is mapped by at most two edges in M_t, one for each endpoint. This establishes the claim that $|\overrightarrow{M} \cap \delta(S_t)| \geq |M_t|/2$. Then, we can conclude that

$$\int_0^\infty \nu(\delta(S_t))dt = \int_0^\infty |M_t|dt \leq 2\int_0^\infty |\overrightarrow{M} \cap \delta(S_t)|dt = 2\sum_{uv \in \overrightarrow{M}} \left(x(u)^2 - x(v)^2\right) \leq 2\nu(\overrightarrow{G_x}).$$

□

Summary and Discussions

Starting from the primal program $\lambda_2^*(G)$ in Definition 8.1, we construct the dual program $\gamma(G)$ in Proposition 8.6 using von-Neumann minimax theorem. Then we use the Johnson-Lindenstrass lemma to reduce an n-dimensional solution to $\gamma(G)$ to a 1-dimensional solution to $\gamma^{(1)}(G)$ in Definition 8.7, where the $\log |V|$ factor in Theorem 8.2 is from this step. Then we consider the matching expansion $\psi_\nu(G)$ in Definition 8.11 as a proxy to the vertex expansion $\psi(G)$ in Definition 7.8, and reduce the Cheeger inequality for vertex expansion in Theorem 8.2 to the Cheeger inequality for matching expansion in Theorem 8.13. The easy direction of Theorem 8.13 can be proved by plugging in a binary solution from matching expansion. Also, there is a good way to understand the easy direction of Theorem 8.2 as a reduction from vertex expansion to the edge conductance of the best reweighted graph. The hard direction of Theorem 8.13 is proved by a truncation step and a threshold rounding step as in the proof for Cheeger's inequality in Theorem 4.3. A key Lemma 8.22 in the hard direction is proved by a combinatorial argument about greedy directed matching. After finding a set of small matching expansion, we can use Problem 8.12 to find a set of small vertex expansion.

Oleskar-Taylor and Zanetti left open the problem of reducing the $\log|V|$ factor in Theorem 8.2 to $\log d$ where d is the maximum degree of the input graph, and the problem of generalizing Theorem 8.2 to arbitrary target probability distribution π. It would also be interesting to construct an example where Theorem 8.2 is nearly tight.

8.4 Problems

Problem 8.23 (λ_∞ and Symmetric Vertex Expansion [BHT00]). *Bobkov, Houdré and Tetali defined an interesting quantity*

$$\lambda_\infty(G) := \min_{x:V\to\mathbb{R},\ x\perp\vec{1}} \frac{\sum_{u\in V} \max_{v:(v,u)\in E}\ (x(u)-x(v))^2}{\sum_{u\in V} x(u)^2}$$

and prove an analog of Cheeger's inequality that

$$\Phi^V(G)^2 \lesssim \lambda_\infty(G) \lesssim \Phi^V(G),$$

where

$$\Phi^V(S) := |V|\cdot\frac{|\partial(S)\cup\partial(V-S)|}{|S|\cdot|V-S|} \quad \textit{and} \quad \Phi^V(G) := \min_{S\subset V}\Phi^V(S)$$

is called the symmetric vertex expansion of the graph. Give a proof of their theorem.

8.5 References

[BDX04] Stephen P. Boyd, Persi Diaconis, and Lin Xiao. Fastest mixing markov chain on a graph. *SIAM Rev.*, 46(4):667–689, 2004. 73

[BHT00] Sergey G. Bobkov, Christian Houdré, and Prasad Tetali. λ_∞ vertex isoperimetry and concentration. *Comb.*, 20(2):153–172, 2000. 83

[LRV13] Anand Louis, Prasad Raghavendra, and Santosh S. Vempala. The complexity of approximating vertex expansion. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 360–369. IEEE Computer Society, 2013. 74

[OZ21] Sam Olesker-Taylor and Luca Zanetti. Geometric bounds on the fastest mixing markov chain. *CoRR*, abs/2111.05816, 2021. 73, 74, 75, 77, 78, 79, 81

[Roc05] Sébastien Roch. Bounding Fastest Mixing. *Electronic Communications in Probability*, 10(none):282 – 296, 2005. 73, 74, 75

[RS10] Prasad Raghavendra and David Steurer. Graph expansion and the unique games conjecture. In *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 755–764. ACM, 2010. 35, 74